

Papers published in *Ocean Science Discussions* are under
open-access review for the journal *Ocean Science*

Forecast and analysis assessment through skill scores

M. Tonani¹, N. Pinardi², C. Fratianni¹, and S. Dobricic¹

¹Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

²University of Bologna, Corso di Scienze Ambientali, Ravenna, Italy

Received: 8 January 2007 – Accepted: 28 January 2007 – Published: 22 February 2007

Correspondence to: M. Tonani (tonani@bo.ingv.it)

OSD

4, 189–212, 2007

**Forecast and
analysis assessment
through skill scores**

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

Abstract

This paper describes a first comprehensive evaluation of the quality of the ten days ocean forecasts produced by the Mediterranean ocean Forecasting System (MFS). Once a week ten days forecasts are produced. The forecast starts on Tuesday at noon and the prediction is released on Wednesday morning with less than 24 hr delay. In this work we have considered 22 ten days forecasts produced from the 16 August 2005 to the 10 January 2006. All the statistical scores have been done for the Mediterranean basin and for 13 regions in which the Mediterranean sea has been subdivided.

The forecast evaluation is given here in terms of root mean square (rms) values. The main skill score is computed as the root mean square of the difference between forecast and analysis (FA) and forecast and persistence (FP), where the persistence is defined as the average of the day of the analysis corresponding to the first day of the forecast. A second skill score (SSP) is defined as the ratio between rms of FA and FP, giving the percentage of accuracy of the forecast with respect to the persistence (Murphy 1993).

The rms of FA is always better than FP and the FP rms error is double than the rms of FA. It is found that in the surface layers the error growth is controlled mainly by the atmospheric forcing inaccuracies while at depth the forecast errors could be due to adjustments of the data assimilation scheme to the data insertion procedure. The predictability limit for our ocean forecast seems to be 5–6 days connected to atmospheric forcing inaccuracies and to the data availability for assimilation.

1 Introduction

The aim of this study is to evaluate the accuracy of the forecast produced at the basin scale from the Mediterranean ocean Forecasting System during the MFSTEP project. The results of this paper should provide the first quantitative information on the accuracy of open ocean forecast system and shows some of the deficiencies of the system.

OSD

4, 189–212, 2007

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

In Murphy (1993) the quality of the forecast is defined as a function of the accuracy that can be estimated comparing the forecast with observations. In this study the reference field will be the analysis produced every week from MFS using an ocean general circulation model (OGCM, Tonani et al., this volume) together with an assimilation scheme (Dobricic et al., this volume). The analysis has been considered as the best estimate of the reality available and the attention will be focused on the performance of the forecast with respect to it.

The literature offers a wide range of possible skills core computations and indices but it has been decided to use a simple rms evaluation following Murphy et. al, (1988) and Demirov et al., (2003) which was done for the previous version of the forecasting system (Pinardi et al, 2003). The study is done only on the forecasted fields of temperature and salinity at different depths and the skill scores are computed over all the basin and in different sub basins characterized by different ocean dynamics and water masses. The period considered goes from August 2005 till January 2006, i.e. 22 weekly ten days forecast cycles. Furthermore we study also the variability of the forecast accuracy due to the seasons, i.e. to the flow field variability.

The paper is organized in the following way. Section 2 describes the MFSTEP forecast system and production chain. Section 3 presents the skill scores and the forecast evaluation. Section 4 discusses the results and Sect. 5 presents the conclusions.

2 Description of MFSTEP forecast system and the forecast production chain

The ocean forecast products are daily mean values of temperature, salinity, three dimensional velocity field and sea level. The forecast production consists of the collection of in situ and satellite data adequately pre-processed, a numerical model and the assimilation scheme. The numerical model is finite differences and it considers an implicit free surface (Tonani et al., this issue): it has been implemented on the Mediterranean Sea with an horizontal resolution of $1/16^\circ \times 1/16^\circ$ of degree and 72 unevenly spaced vertical levels and it is forced with ECMWF analysis and forecast atmospheric field.

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The assimilation scheme used is a reduced order Optimal Interpolation system implemented in the Mediterranean Sea at different levels of complexity for the past five years (Dobricic et al. 2006 and this volume).

5 The assimilated data are temperature and salinity vertical profiles from XBT and ARGO and sea level anomalies from altimetry. The data are collected and prepared every week (see Appendix A) on Tuesday and the system is run during the night between Tuesday and Wednesday. The system produces the analyses for the previous fifteen days and ten days of forecast, as shown in Fig. 1. All the data sets are assimilated once in a single analysis and using the same background error covariance matrix.
10 This is a major differences respect the system used in the previous operational system described by Demirov et al., (2003). To produce an analysis, the OGCM makes a 24 h simulation, misfits (differences between observations and model at the same space and time location) are calculated each day using the FGAT (First Guess at Appropriate Time) method. The analysis is computed at the end of each day and then the model
15 is re-started from this new initial condition. The assimilation of SLA observations uses an estimate for the mean dynamic topography (Rio et al., 2004) that considers observations and model data. Satellite data of Sea Surface Temperature are used to correct the surface heat fluxes of the model through a relaxation correction formula (Pinardi et al., 2003)

20 The MFSTEP forecast is produced at a start time J which is Tuesday noon every week. The preparation and running of the forecast is done through an automatic procedure which has been set up and tested during the MFSTEP project. The operational chain is activated as soon as the ECMWF forcing data, the daily satellite SST and the SLA along track data are available every tuesdays. The procedure of data acquisition and preparation is described in Appendix A. The 10 days forecasts fields and the
25 last seven days of analysis are disseminated through an ftp site as soon as they are produced.

After the forecast has been carried out, the post processing procedures begin, both in terms of skill scores and end-users services. Usually all the procedures finish in

**Forecast and
analysis assessment
through skill scores**

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

the morning of Wednesday, that means that there is a delay of less than 24 h in the forecast release. This is a good improvement respect the system that was running in the previous version of the operational system (Demirov et al., 2003; Pinardi et al. 2003), where the forecast was released with three days delay. This improvement is mainly due to the release of the SST products and to the automatization of the forecasting procedures.

3 Data and methods

The data used in this study are the ten daily mean forecasts produced every week from the 16 August 2005 to the 10 January 2006, that means 22 weeks of forecasts.

The daily mean values of the analyses and forecasts are then compared in order to quantify the forecast skill scores (Murphy, 1988, 1993). The differences between analysis and forecast are due to:

- assimilation of sst, sla, xbt and argo data that are used for the analysis production;
- the atmospheric forcing that in the case of the analysis is deduced from ECMWF surface analysis fields while in the ocean forecast case is calculated from the atmospheric surface field forecasts.

The 10 days of forecast are compared with the corresponding analysis which have been produced two weeks after the forecast production day. All the computations in this study are done using the best available analyses for each considered day.

The root mean square of the forecast vs. analysis (FA) is computed as follow:

$$\text{rms}(t) = \sqrt{\frac{\sum_{1}^N (X_f(t) - X_a(t))^2}{N}} \quad (1)$$

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

where $X_f(t)$ is the daily mean of the temperature or salinity field from the forecast for the day t at a selected depth, while $X_a(t)$ is the daily mean of the temperature or salinity field from the analysis for the same day and depth. N is the total number of ocean point at the selected depth.

5 The root mean square of the forecast vs. persistence (FP) is computed as follow:

$$\text{rms}(t) = \sqrt{\frac{\sum_{t=1}^N (X_f(t) - X_a(t=1))^2}{N}} \quad (2)$$

where the persistence is considered to be the analysis daily mean field at $t=1$. To be noticed is that the forecast is initialized with an instantaneous field so that our persistence is not really equal to what is used in the atmospheric context. However, since
10 the time scales of the ocean are slower we believe that this could be a good estimate of the persistence skill score.

It is important to point out that the assumption that the analysis is the best estimate of the reality is due to the extremely non-homogeneous observational coverage. The XBT, ARGO and altimetry tracks are shown in Fig. 2 to provide insight in the horizontal
15 sampling coverage.

The spatial distribution of the in situ observations is very uneven but nevertheless they are shown to be fundamental to correct for model inaccuracies (Demirov et al., 2003, Dobricic et al., 2006). The rms of misfits is shown also in Fig. 2, indicating that the rms of the misfit decays with time at all levels, showing a beneficial impact of the
20 data on the model accuracy. Thus the analyses are a reasonable good estimate of the reality and we will use them to evaluate the forecast performance.

In this study we have computed the rms of the FP and FA at selected depths of 5 m, 30 m, 150 m, 300 m and 600 m. The rms have been computed for each of the 22 ten days forecasts and then averaged in time for the whole period duration or by month for
25 September, October, November or December 2005.

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Following Murphy et al., (1988) the Percentage Skills Score (SSP) is computed as:

$$SSP = \left(1 - \frac{FA}{FP}\right) * 100 \quad (3)$$

SSP has been computed for both temperature and salinity at the depths of 5, 30, 150, 300 and 600 m of model depth. SSP is equal 100% if FA is equal zero, that means a perfect forecast. Otherwise SSP is equal 0 if FA is equal FP. If the accuracy of the forecast is greater than persistence then $SSP > 0$ and vice versa $SSP < 0$ if the forecast is less accurate than persistence.

4 Forecast evaluation results

4.1 Rms of of FA and FP for the entire basin

The rms of FP and FA for the whole Mediterranean Sea is shown Fig. 4. The rms values of FP are always bigger than the rms of FA and they increase rapidly with time especially at the depths of 5 and 30 m. However, the rms of FP shows some saturation at the end of the 10 days forecast period while the rms of FA seems to be still growing rapidly. At the tenth day the rms of temperature FP reaches a value that is twice the equivalent for FA and one third larger for the salinity. The values of rms FA and FP at 150 m are small and equivalent especially for the temperature. For the salinity the rms of FA is better than for FP but again much smaller than in the surface layers. The simple rms of FA and FP cannot show the differences between analysis and forecast at this depth due to a combination to effects: 1) the ocean variability is smaller at depth than at the surface and ten days are short to show differences; 2) the averaging over the whole basin makes the analysis very little different from the forecast since of the large data void regions.

The surface is strongly influenced by the quality of the atmospheric forcing used that is one of the main differences between analyses and forecasts. Figure 3 shows the

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

standardized rms of the difference between analysis and forecast from ECMWF for all the surface atmospheric fields used by the ocean model. We used practically Eq. (1) divided by the standard deviation of the rms of each atmospheric field. It is clear from the figure that the degradation of the atmospheric forcing has a linear behaviour. That means that the error due to the atmospheric forcing constantly increases in the ten days of the forecast that is what is found also in Fig. 4 for the rms of FA. This explains why the rms of FA does not saturate in time but follows the main source of errors that is connected to the atmospheric forcing inaccuracies.

Figure 5 shows the SSP for each 10 days of the forecast averaged over the 22 weeks, for temperature and salinity at five depths. At the first day of the forecast, SSP is always equal zero due to its definition (see Eq. 3). SSP is positive for salinity and temperature at 5 and 30 m, reaching maximum values of 50% and 40% respectively at the 5–6 day of forecast. After that day the values start to slightly decrease maintaining values of 45–35%. This means that the forecast improves our estimate with respect to persistence for the first 5–6 days of the forecast.

For the upper levels the behaviour of the SSP is approximately the same for temperature and salinity. However, it is also evident from Fig. 5 that while at the surface the forecast accuracy is better than persistence in the first 2–3 days, down to 150–300–600 m the persistence can beat the forecast and only after the 3–4 day the forecast gains accuracy with respect to persistence. We argue that this is most probably related to an ocean adjustment process introduced by the data assimilation which could deteriorate the model results for the first days. The investigation of the reasons for this behaviour is outside the scope of this paper and it is an active line of research for the development of new data assimilation schemes.

4.2 Rms of FA and FP in different Mediterranean regions

Rms of FP and FA have now been computed for 13 different regions, shown in Fig. 6, in order to better understand how the quality of the forecast varies in the different areas of the Mediterranean sea. The subdivision of the Mediterranean Sea into this 13

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

regions have been done taking into account the water mass properties and the different dynamics. The SSP skill scores have been computed as averages for each of the 13 regions and it is likely that the results from region to region could be different due to several factors as: 1) the quality of the analysis which depends on the data quantity in each region; 2) the different dynamics of each region in terms of advection and mixing of the thermocline which is a dominant process in the months considered and that makes the impact of data insertion different.

As shown in Fig. 2 the data distribution is sparse that means that in some region the assimilation of data could be really efficient and in some other in completely inefficient because the number of data is equal zero. Regions like the Alboran Sea (region n. 1) and the South-West Ionian (region n. 7) where there are no in situ data and no SLA data will have high values of SSP simply due to the fact that the accuracy of the analysis is low. In these regions the differences between forecast and analysis are small and only due to the differences in the atmospheric forcing and to the assimilation of the satellite SST.

The SSP for some of the regions is shown in Fig. 7. Regions 1 and 7, as described above, have the highest values of the SSP for both salinity and temperature at all depths. This is due to the low accuracy of the analysis, as described above.

Regions like the Algerian Basin (region n. 2) where only SLA data are available have positive SSP values that are lower than in regions 1 and 7 and with values closer to the basin mean of Fig. 5. In region 2 the temperature has better score at the surface than the salinity and this could be due to the high variability of the salinity in this area characterized by the inflow of relatively fresh water from the Atlantic Ocean through the Strait of Gibraltar.

In the Gulf of Lion area (region n. 3) where there is a good coverage of data, both satellite and in situ, SSP have relatively high value, around 40% for the temperature at the depths of 5 and 30 m and 30% for salinity. SSP has negative values below the depth of 150 m for both temperature and salinity fields and at 600 m negative values persist till the 6 day of forecast. This is a longer time than in Fig. 5 where it was negative only for

Forecast and analysis assessment through skill scores

M. Tonani et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

the first four days of the forecast. This means that the impact of the data assimilation in this region determines an adjustment time longer than in the overall basin.

A similar situation happens also in the North Levantine Basin (region n. 11) where the temperature field at 5 m has good SSP reaching values of 40% and the salinity of 20% but the temperature at 150 m and the salinity and temperature at 300–600 m show negative SSP. In this region, the thermocline is approximately located between 50 and 150 m and the data assimilation corrections inserted in this region could cause large adjustments, causing the SSP to be low and for several days.

The south-eastern Levantine basin (region 13) has values of SSP relatively high at all depths and only in the first two days of forecast SSP is negative for both temperature and salinity at 600 m.

4.3 Monthly mean skill scores for the entire basin

The monthly mean rms of FA and FP and the SSP for the months of September, October, November and December 2005 have been computed at 5 m and 30 m for the entire Mediterranean basin. Table 1 shows the data availability during the different months and in Fig. 8 we show the skill scores. The rms of FA and FP shows that the 10 days rms values are higher in Sept–Oct than in the Nov–Dec for both temperature and salinity. This is probably due to the fact that the continuous insertion of data brings the model toward a better forecast with time as shown also in Fig. 2 as decreasing rms of misfits.

The FP and FA curves for the temperature show very well the monthly decreasing trend variation. In addition, the FP curves show a steepness increase going from summer to winter. This could be related to the increased time variability of the ocean dynamics that makes persistence a worse forecast going from summer to winter. The rms FA and FP for salinity show a reduction of the FA error by a factor of 3 between September and December, much more than for temperature. Thus the impact of assimilation with time and its beneficial impact on the forecast skill scores is higher for salinity than for temperature.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The lower panels of Fig. 8 show the values of SSP at 30 m for the considered months. In the case of the temperature it is clear how the months of November and December have the highest SSP values which increase quite rapidly (60%) in the first 4 days of forecast reaching the maximum value of 62%. In September and October SSP reaches a maximum value of 35–40%, half of the value achieved in November and December. This could be related to the quality of the increased quality of the analysis as shown in Fig. 2.

In all the curves, after the 5 forecast day the SSP curves flatten out and asymptote to a fixed value. This means that the forecast quality saturates and does not continue to grow as instead it seems to be the case from rms FA score curves (also evident in Fig. 3). The effect of increasing error in the atmospheric forcing which causes the rms of FA to grow is cancel out by the same effect in the rms FP values, so that the SSP is a more convenient score to be used in order to judge the loss of predictability of the forecast. We can then say that the predictability saturates around 5–6 days for the high resolution ocean forecasting system set up here.

Another important feature of Fig. 8 is the negative SSP values only for the October case for both salinity and temperature. It is very well known that in the Mediterranean Sea this month coincides with the largest thermocline gradients and the maximum model errors due to the inaccuracies of the vertical mixing schemes and the air-sea interactions parametrizations (Pinardi et al., 2003). We argue then that the negative SSP score in this month is due to model errors that limit the capability of the model to reproduce the shape of the thermocline.

In conclusions, the major differences from month to month are probably due to a combination of data assimilation adjustment and model inaccuracies that depends on the dominant ocean dynamical regimes in the different seasons. Even if the period of study is not sufficiently long to properly study the variability of the forecast accuracy due to the seasons we believe this is a robust result and it will characterise the forecast errors in all the strongly seasonal basins at mid-latitudes.

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

5 Conclusions

This work is the first attempt to estimate the quality of the forecast produced during the MFSTEP project and it defines a forecast evaluation protocol.

5 The forecast has been evaluated in terms of rms difference between the forecast and the analysis (FA) and the forecast and the persistence (FP). A new skill score has also been defined, SSP, that is shown to be extremely valuable to understand possible cause of the predictability limits in the ocean.

10 The accuracy of the MFSTEP forecast for the period considered is always, except down to 600 m, better than persistence and with values of SSP as high as 60%. However, care should be taken at depth, where the data insertion might cause adjustments that deteriorate the forecast quality with respect to persistence.

15 Choosing to use the analysis as the reference truth, its quality or conversely the data scarcity, determines the skill score values. A large SSP value could be simply due to the missing data and thus a low accuracy of the analysis instead of a high accuracy of the forecast.

20 The predictability limit of temperature and salinity has been found to be 5 days for a high resolution forecasting system that uses uncoupled atmospheric forecast forcing. This might be a little underestimation of the real predictability limit in the ocean and new skill scores should be developed such as anomaly correlations that will not weight so much the phase shift error.

Appendix A

The operational procedure

25 The data collected every week on Tuesday consider the oceanic observations in near real time (NRT) and the meteorological fields from the European Middle range Weather Forecast (ECMWF). The NRT oceanic observations include:

OSD

4, 189–212, 2007

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

- Sea Surface Temperature (SST)
- Sea Level Anomaly (SLA)
- Vertical temperature profiles from eXpandable BathyThermograph (XBT) collected by Voluntary Observing Ship (VOS)
- 5 – Vertical temperature and salinity profiles observed by ARGO floats deployed by MFSTEP and NAVOCEANO.

10 The SST product is based on night-time AVHRR images acquired and processed by Centre de Meteorologie Spatiale (CMS) and Gruppo di Oceanografia da Satellite, Institute of Atmospheric Sciences and Climate of the Italian National Research Council (GOS-CNR-ISAC). In particular, mapped AVHRR SST acquired and processed at CMS are uploaded from CMS to GOS -CNR-ISAC, and merged to the AVHRR SST acquired and processed at GOS-CNR-ISAC. The data consist of nighttime values interpolated on the model grid.

15 The SLA data, that are used in the MFS assimilation scheme, are collected by Collection and Localization Satellitaire (CLS) in Toulouse. The product consists of along track SLA values from two satellite missions: Jason1 and Geosat Follow On (GFO). The products delivered are computed with precise Orbit Error Reduction and Long Wavelength Error reduction (OER), (Buongiorno Nardelli et al., 2003).

20 The vertical temperature profiles are collected along several routes in the Mediterranean Sea by VOS (Manzella et al. 2003), The near real time quality control is done by ENEA and the procedure is repeated at INGV before the data are assimilated and interpolated on the vertical model grid:

25 The vertical temperature and salinity profiles collected by ARGO floats are retrieved once a week from the Coriolis data centre in IFREMER (Brest) On these data a first quality control is done by IFREMER then a second quality control is done by INGV and it contains:

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Check of the date and position flag: the profile is rejected if the flag for both is different from 1;
- Check of the flag of each T and S value: couples of T and S, corresponding to a pressure level, are rejected if the flags are different from 1. Check on the pressure flag is not done;
- Elimination of undef values;
- Calculation of the Brunt Vaisala frequency: values of T and S corresponding to negative frequencies are rejected;
- Rejection of the whole profile if the distance T and S in the first 150 m is greater than 40 m;
- Organization of the data in ascii files containing all the data collected in the same day. Then before are used in the assimilation scheme, the data are interpolated on the model levels.

The analyses and forecast of the meteorological fields from ECMWF are provided each day by Meteo France. The ECMWF atmospheric forcing is constituted of 6 fields, the mean sea level pressure, zonal and meridional component of the wind at 2 m, total cloud cover, temperature at 2 m, dew point temperature at 2 m. The time step for analyses fields is 3 h while for forecast fields is 3 h until 72 h and 6 h after. They are on a regular 0.5×0.5 grid on approximately the same sub domain of MFS (30 N–48 N–19 W–42 E).

References

Buongiorno Nardelli, B., Larnicol, G., D'Acunzio, E., Santoleri, R., Marullo, S., and Le Traon, P.-Y.: Near Real time SLA and SST product during 2-years of MFS pilot project: Processing, analysis of the variability and of the couplet patterns, *Ann. Geophysicae*, vol.21, n.1, 189–204, 2003.

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Demirov, E. and Pinardi, N.: Simulation of the Mediterranean Sea circulation from 1979 to 1993: Part I. The interannual variability, *J. Mar. Syst.*, 33–34, 23–50, 2002.
- Dobricic, S., Pinardi, N., Adani, M., Bonazzi, A., Fratianni, C., and Tonani, M.: Mediterranean Forecasting System: An improved assimilation scheme for sea-level anomaly and its validation, *Q. J. R. Meteor. Soc.*, 131, 3627–3642, 2005.
- Dobricic, S., Pinardi, N., Adani, M., Tonani, M., Fratianni, C., Bonazzi, A., and Fernandez, V.: Daily oceanographic analyses by the Mediterranean basin scale assimilation system, *Ocean Sci. Discuss.*, 3, 1977–1998, 2006, <http://www.ocean-sci-discuss.net/3/1977/2006/>.
- Murphy, A. H.: Skill score based on the mean square error and their relation to the correlation coefficient, *Monthly Weather Review*, 116, 2417–2424, 1988.
- Murphy, A. H.: What is a good forecast? An essay on the nature of goodness in weather forecasting, *Weather Forecasting*, 8, 281–293, 1993.
- Pinardi, N., Allen I., Demirov E., De Mey, P., Korres, G., Lascaratos, A., Le Traon, P.Y., Maillard, C., and Manzella, G. M.: The Mediterranean ocean Forecasting System: First Phase of implementation (1998–2000), *Ann. Geophysicae*, vol.21, n.1, 189–204, 2003.
- Demirov, E., Pinardi, N., Fratianni, C., Tonani, M., Giacomelli, L., and De Mey, P.: Assimilation scheme of the Mediterranean Forecasting System: operational implementation, *Ann. Geophysicae*, vol.21, n.1, 3–20, 2003.
- Tonani, M., Pinardi, N., Dobricic, S., Pujol, I., and Fratianni, C.: A high resolution free surface model on the Mediterranean Sea, accepted, *Ocean Sci. Discuss.*, 2007.
- Manzella, G. M., Scoccimarro, E., Pinardi, N., and Tonani, M.: Improved near-real time data management procedures for the Mediterranean ocean Forecasting System-Voluntary Observing Ship program, *Ann. Geophysicae*, vol.21, n.1, 49–62, 2003.

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Number of satellite data from SLA and in situ temperature and salinity profiles by XBT and ARGO collected during the months of September, October, November and December 2005. The estimate of the in situ data (as for Table 2) has been done referring to the depth of 30 m.

		SEP	OCT	NOV	DEC
REG. 1	SLA	19	16	39	11
	S	0	0	0	0
	T	0	0	0	0
REG. 2	SLA	488	752	744	834
	S	8	5	8	7
	T	8	5	8	7
REG. 3	SLA	455	679	703	712
	S	16	19	13	12
	T	16	19	13	12
REG. 4	SLA	208	305	375	364
	S	11	9	8	8
	T	11	9	8	8
REG. 5	SLA	0	0	0	0
	S	0	0	0	0
	T	0	0	0	0
REG. 6	SLA	276	353	427	346
	S	10	5	9	7
	T	10	5	9	7
REG. 7	SLA	0	2	2	2
	S	0	0	0	0
	T	0	0	0	0
REG. 8	SLA	634	1021	965	942
	S	14	15	16	13
	T	14	15	16	13
REG. 9	SLA	29	69	62	68
	S	0	0	0	0
	T	0	0	0	0
REG. 10	SLA	224	423	385	395
	S	0	0	3	14
	T	0	0	3	14
REG. 11	SLA	371	572	508	521
	S	22	26	24	14
	T	22	26	24	14
REG. 12	SLA	269	434	415	383
	S	0	6	8	22
	T	0	6	8	22
REG. 13	SLA	151	272	248	272
	S	11	6	12	13
	T	11	6	12	13

Forecast and analysis assessment through skill scores

M. Tonani et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Forecast and analysis assessment through skill scores

M. Tonani et al.

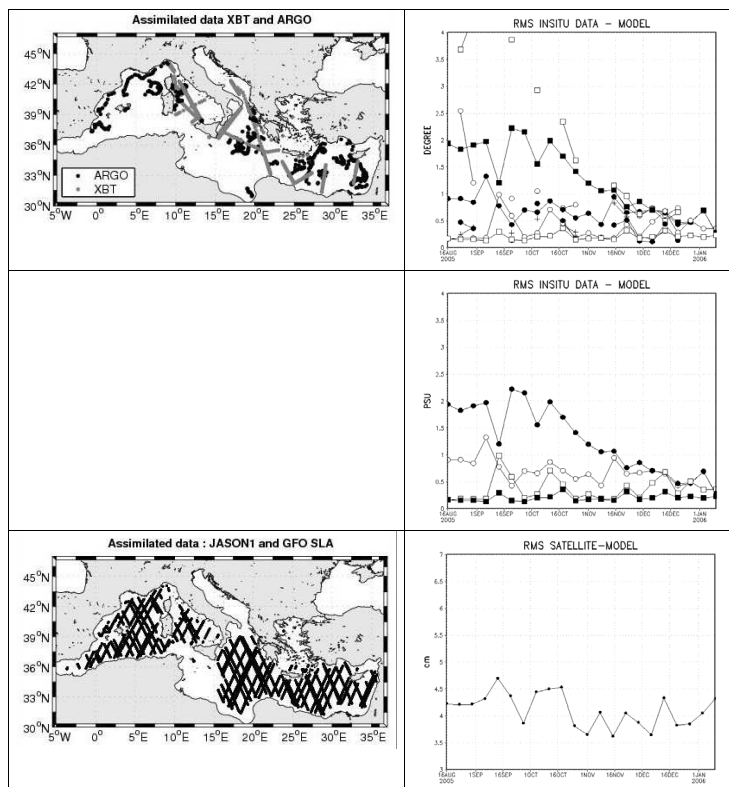


Fig. 2. Maps of data (satellite and in situ) collected over all the period of study. The rms between model and observations (misfit) is shown, computed as averages for the whole Mediterranean Sea. The temperature and salinity rms have been computed at the depths of 5, 30, 150 and 300 m, for the temperature they are computed for both XBT and ARGO data, for the salinity from the ARGO.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Forecast and analysis assessment through skill scores

M. Tonani et al.

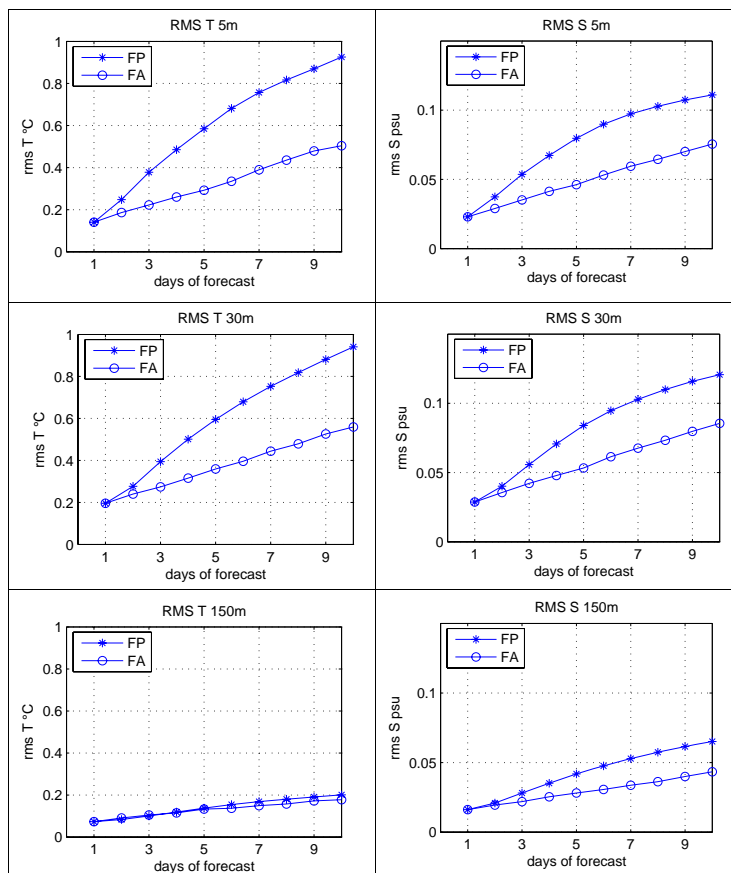


Fig. 3. Rms of FA and FP for salinity and temperature computed at 5, 30 and 150 m. The values are means over the 22 ten days forecasts and over the whole basin.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Forecast and analysis assessment through skill scores

M. Tonani et al.

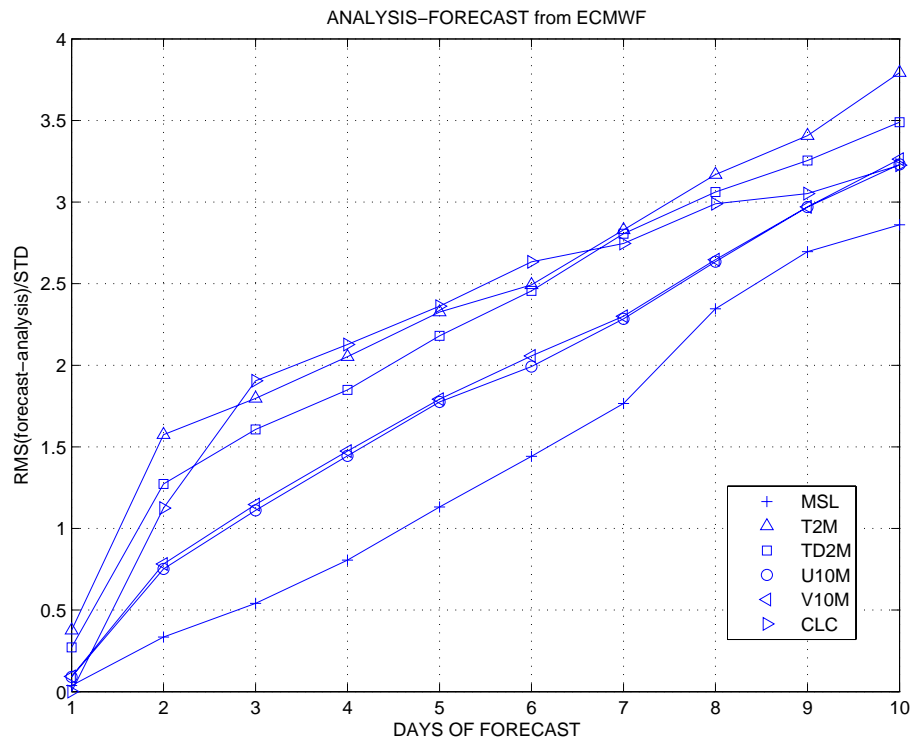


Fig. 4. Rms between ECMWF analyses and forecast atmospheric fields, mean value over the 22 ten days forecasts considered in this work. The rms has been normalized by the std of the rms of each atmospheric field considered.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

Forecast and analysis assessment through skill scores

M. Tonani et al.

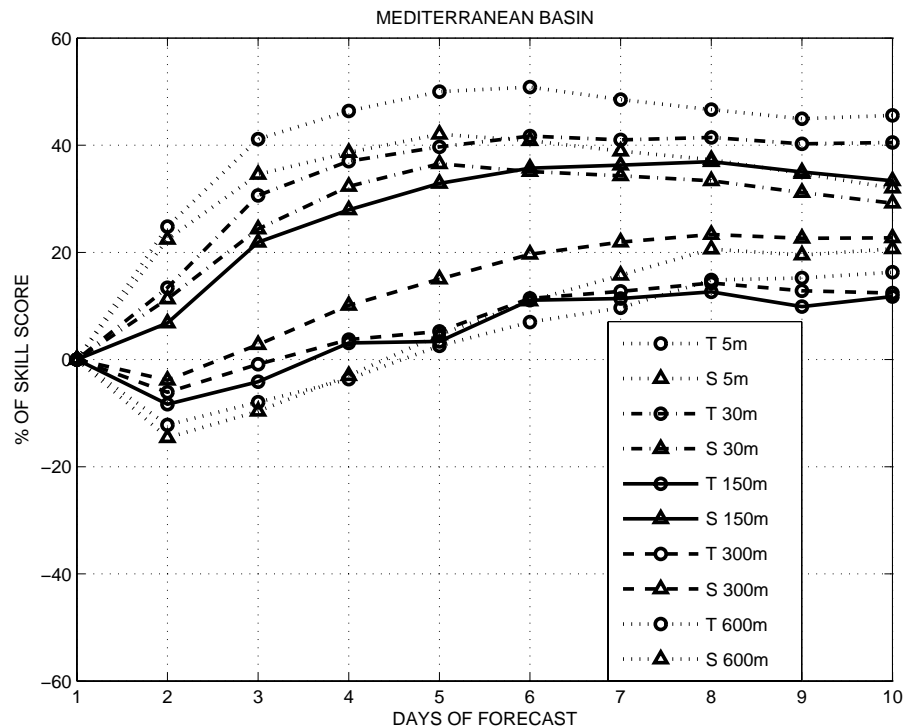


Fig. 5. Skill Score (SSP) computed at 5, 30, 150, 300 and 600 m of depth for the salinity and the temperature fields. The values are computed as means over the 22 ten days forecasts and over the whole the Mediterranean basin.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Forecast and
analysis assessment
through skill scores**

M. Tonani et al.

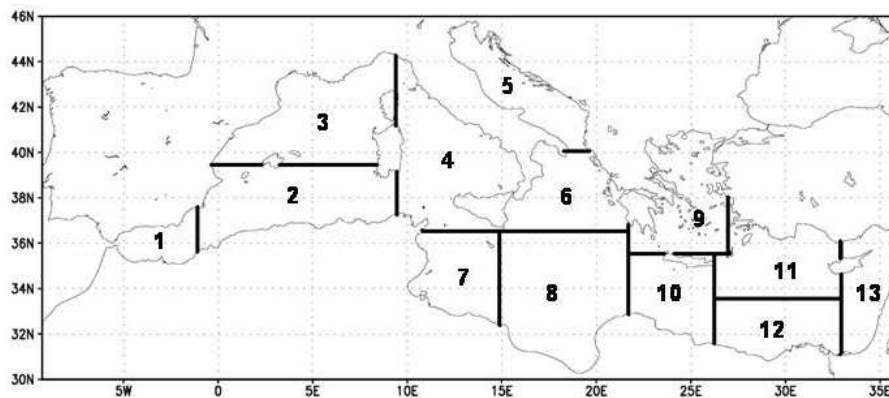


Fig. 6. The Mediterranean Sea regions where statistics has been computed.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Forecast and analysis assessment through skill scores

M. Tonani et al.

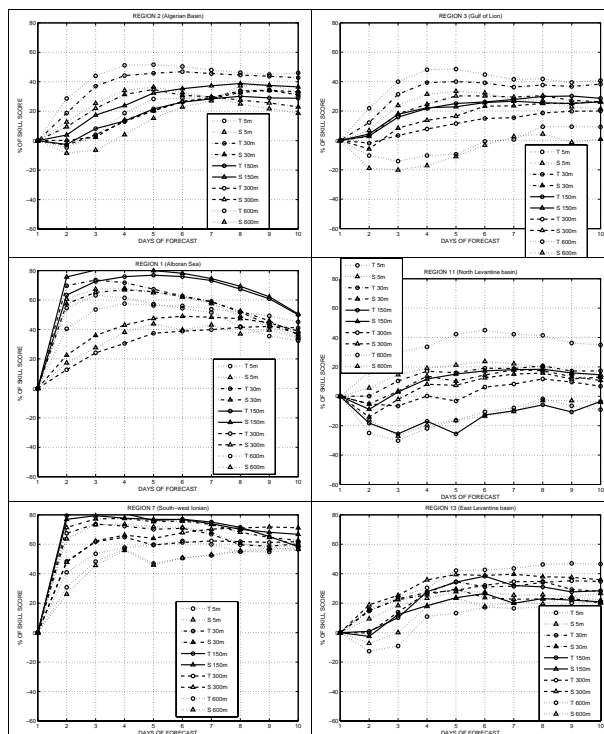


Fig. 7. Skill Score (SSP) computed at 5, 30, 150, 300 and 600 m of depth for the salinity and the temperature fields. The panels (from the up-right corner in clock wise direction) show respectively the values computed over region 2 (Algerian basin), region 3 (Gulf of Lion), region 1 (Alboran Sea), region 11 (north Aegean Sea), region 7 (South-west Ionian sea) and region 13 (South-eastern Levantine Sea).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Forecast and analysis assessment through skill scores

M. Tonani et al.

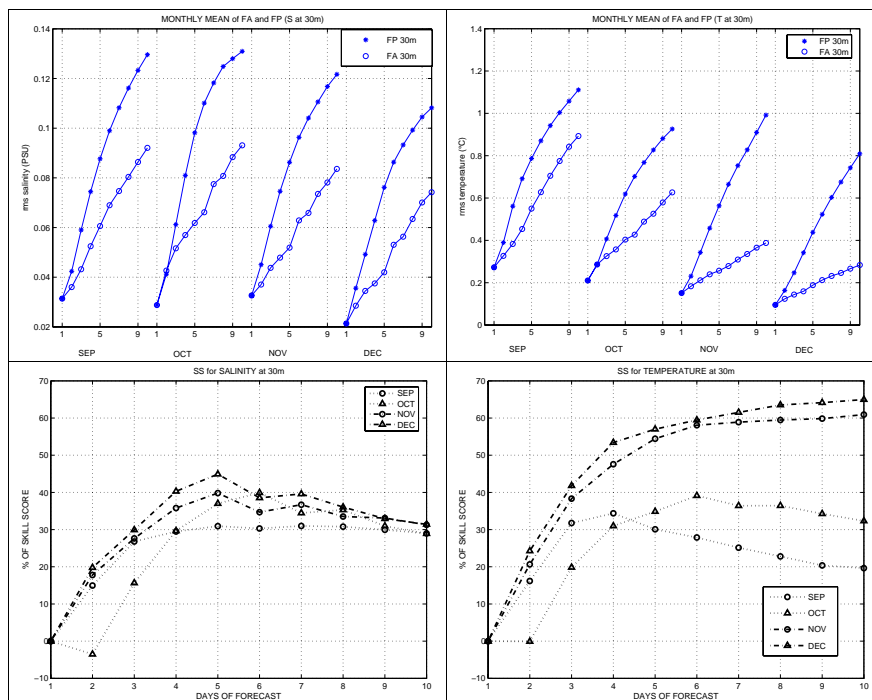


Fig. 8. Monthly mean of rms and SSP computed for all the Mediterranean basin at the depths of 30 m for the salinity and the temperature fields.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion